Firm size distribution and exporting behaviour: an empirical analysis of power-law behaviour of turkish firms

Cansin Pek

FIRM SIZE DISTRIBUTION AND EXPORTING BEHAVIOUR : AN EMPIRICAL ANALYSIS OF POWER-LAW BEHAVIOUR OF TURKISH FIRMS

Author:

Cansin Pek

Supervisor:

M. Angelo Secchi

Professor of Economics

Universite Paris 1 - UFR 02 Sciences Economiques - Master 2 Recherche Economie Appliquée

June, 2012

L'Université de Paris 1 Panthéon Sorbonne n'entend donner aucune approbation, ni désapprobation aux opinions émises dans ce mémoire; elles doivent être considérées comme propre à leur auteur.

Abstract

A general equilibrium model of international trade with heterogeneous firms, under the assumption that the distribution of productivity across firms is Pareto, delivers systematically different power law exponents for exporting and non-exporting firms. In this setup, the presence of international trade systematically affects the firm size distribution to make it more heavy-tailed. This model predicts that the power law exponent for exporters should be strictly lower in absolute value than the power law exponent for non-exporters. Following the propositions made in the literature, we estimate the power law exponent for a large sample of Turkish firms. We also question the applicability of the OLS regression in the context of power law estimation and provide maximum likelihood estimates, which have been proven to be consistent and efficient in this context. Along with the maximum likelihood estimates, we also provide the CDF and $\ln(Rank-1/2)-\ln(size)$ estimates. Our results provide supporting evidence for the theoretical predictions, according to which the distribution of firm size has heavier tails due to exporting behaviour.

1 Introduction

Power law distributions appear in a diverse range of natural and man-made phenomena. Firm sizes [3; 11], population of agglomerations [24], the frequency of occurence of unique words in any human language, the sales volume of bestseller books, the intensities of solar flares, personal annual incomes [5] among many other phenomena have been considered to follow power law distributions, taking the form $Pr(X > x) = Cx^{-\alpha}$, where $Pr(X > x)$ is the probability that a random variable X is greater than x, α the scaling parameter and C some constant.

The literature on firm size distribution have shown that, the probability density is described by a straight line on a log-log scale, with a slope of (approximately) -2, i.e. a power law distribution with the scaling parameter equal to 1 [3][1], [11][2]. However, the deviations from the power law form occur for very small and very large sizes. On closer inspection, the straight line mentioned above is actually somewhat concave. It is also reported that the empirical density can be described quite well by a log-normal density function[15; 14; 21]. However, Quandt (1966) [22] and Sutton (1997) [20] both argue that the empirical density cannot be decently described either by a power law nor by a log-normal.

The empirical literature on firm size distribution taking into account the impact of international trade is not very vast. Eaton et al. (2011) [8] use a detailed data on the exports of French manufacturing firms. They use a variant of Melitz-Chaney model to explain the systematic regularities that

[1] [3], using Census data, finds a power law with exponent $\alpha = 1.059 \pm 0.054$.

[2] Using a representative sample of French firms, [11] find an α of 1.017 ± 0.032.

they observe in the data. Core elements of their model are that firm productivity distribution is Pareto (as in [16; 6]), demand is Dixit-Stiglitz, and markets are separated by iceberg trade barriers and require a fixed cost of entry. In order to overcome the discrepancies between the empirical facts and the model, they introduce market and firm-specific heterogeneity in entry costs and demand and an increasing marketing cost to reach additional consumers in each country as in Arkolasis (2008) [1]. One of their findings is that the size distribution of exporters in a given foreign market is Pareto, after a certain minimum threshold. They also find that the most productive firms serve markets with higher fixed costs of exporting. Eaton et al. (2011) [8] use the method of simulated moments and find an α of 1.46 ± 0.10.

Arkolasis (2008) develops a dynamic trade model with product differentiation, heterogeneous productivity firms, and increasing marginal market penetration costs. Assuming that there is a continuous entry of firms at a certain rate and productivities of entrants evolve according to a geometric Brownian motion, the model endogenously generates a right tail cross-sectional Pareto distribution of firms' productivities. The cross-sectional predictions of his model for the distribution of domestic and exporting sales of firms are consistent with firm-level data [1].

di Giovanni et al. (2011) [11] estimate the power law coefficient for a large sample of French firms, taking explicitly into account the impact of international trade. Using the Melitz (2003) model and assuming the distribution of firm productivities is Pareto, they show that the power law exponent for exporting firms is different and lower (in absolute value) than that for non-exporting firms. Using linear regression methods, they find estimates of the power law exponent for exporting firms, which are systematically lower than that for non-exporting firms ($\alpha = 0.94 \pm 0.042$ for exporters and $\alpha = 1.055 \pm 0.011$ for non-exporters.). In order to obtain reliable estimates for α, they propose two methods: first one consists in estimating the power law exponent for non-exporters only and second one, estimating α using the domestic sales data for all firms. This is the approach followed in this paper.

The precise value of the scaling parameter matters in various contexts. For example, Gabaix [9] argues that many economic fluctuations are not due to small diffuse shocks that directly affect every firm; instead idiosyncratic shocks to large firms potentially generate nontrivial aggregate shocks that affect macroeconomic variables. He shows that if the firm size distribution has thin tails, i.e. $\alpha>2$, then the GDP volatility decays according to $1/\sqrt{N}$ (where N is the number of firms in the economy); however, if the firm size distribution has fat-tails, i.e. $\alpha<2$, then the GDP volatility decays as $1/N^{1-\frac{1}{\alpha}}$, much more slowly than $1/\sqrt{N}$. This means that, when the firm size distribution is fat-tailed, idiosyncratic shocks to individual firms will have a nontrivial aggregate effect, opposing the simple diversification argument where aggregate fluctuations should have a size proportional to $1/\sqrt{N}$. di Giovanni and Levchenko (2009) [12] takes the issue further and finds that openness to international trade can have an impact on aggregate fluctuations by increasing the relative importance of large firms and thus making the economy more granular.[3] Another reason why empirical power law estimates are important is that they can be used to pin down crucial parameters (such the elasticity of substitution between goods) in calibrated

[3]An economy is said to be granular, that is the idiosyncratic shocks to firms result in aggregate fluctuations, if the distribution of firm size follows a power law with a scaling parameter sufficiently close to 1 in absolute value.

heterogeneous firms models and that quantitative results often depend very sharply on the precise parameter values that govern the distribution of firm size [11].

In the literature, the most frequently used method for estimating the power law coefficient (i.e. the scaling parameter) is linear regression. Because the power law takes the form $Pr(X > x) = Cx^{-\alpha}$, taking the log of both sides gives $\ln Pr(X > x) = -\alpha \ln x - \ln C$, which is a straight line, with a slope of α. A line is then fitted to the data by OLS to find the scaling parameter. However, it is argued that this approch has some drawbacks. First, because, some of the assumptions of the OLS method are not fulfilled when used for estimating a power law model. Second, this approach does not allow us to test formally the plausibility of the power law for the data in question. Here, we will use the maximum likelihood estimation and test the plausibility of the power law hypothesis in order to get consistent estimates, following the methods proposed by [7], thus avoiding the drawbacks of the OLS method. Merely for the sake of comparison of our empirical findings with those of other studies, we have nonetheless recourse to two linear regression methods.

In this paper, we aim to provide accurate estimates of power law parameters for Turkish firms, taking into account the exporting behaviour, implementing a principled statistical framework to quantify and test the power law model in empirical data, following [7; 11]. The paper is organized as follows. In Section 2, we present the theoretical framework, and show how the scaling parameter is affected when international trade is taken into account. In Section 3, we analyze the linear regression method in power law estimation and highlight its limitations; then describe the estimation (MLE) and testing procedure of the power law hypothesis. In Section 4, we describe the data and present the empirical methodology. In Section 5, the results are presented. We conclude in Section 6.

2 Theoretical Framework

In this section, we will describe the analytical power law relationship in the distribution of firm size, employing the Melitz (2003) model [18][4]. We assume that the distribution of firm productivities is Pareto. Melitz (2003) [18] develops a dynamic industry model with heterogeneous firms to analyze the intra-industry effects of international trade. He adds firm level heterogeneity in productivity to the classical framework of Krugman (1980), assuming that firms differ in terms of marginal productivity of labour; that the productivity of each firm is randomly drawn from a distribution and that firms do not know their productivity prior to starting production. He also assumes an iceberg type of variable cost of exporting. In this setup, he shows that exposure to trade leads to the entry of the more productive firms in the export market (because of the presence of a fixed cost of entering foreign markets, only a subsection of firms will be able to export), while the less productive firms continue to serve the domestic market and the least productive firms exit. In this model, the more productive the firm, the more markets it serves and has larger size (earns larger profits), thus we expect the firm size distribution to have a heavier upper tail and thus, a Power law exponent that is lower in absolute value, when compared to the non-exporting firms.

[4]This section draws heavily on [11]

Demand

A continuum of firms produces a unique CES variety, using the labour factor, the only factor of production. Given the mass of varieties J_n supplied to the market, the consumer's utility function in market n is

$$max \left[\int_{J_n} c_{ni}^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

and the income constraint is,

$$\int_{J_n} p_{ni} c_{ni} di = Y_n,$$

where c_{ni} denotes the consumption of good i in country n, p_{ni} is the price of good i in country n, Y_n is total expenditure in the economy, and J_n is the mass of varieties supplied in country n. These goods are substitutes, which imply $0<\frac{\varepsilon-1}{\varepsilon}<1$ and an elasticity of substitution between any two goods of $\varepsilon>1$. By maximizing the utility subject to the income constraint, we can calculate the market demand for the variety i as a function of its price:

$$c_{ni} = \frac{Y_n}{P_n^{1-\varepsilon}} p_{ni}^{-\varepsilon}, \tag{1}$$

where P_n denotes the ideal price index in economy n, $P_n = \left[\int_{J_n} p_{ni}^{1-\varepsilon} di \right]^{\frac{1}{1-\varepsilon}}$.

Supply

For each country, there is a pool of prospective entrants into the industry, which is denoted by $\bar{I}_n$. Each potential entrant produces a unique CES variety. There are both fixed and variable costs of production and trade. A firm $i \in [0, \bar{I}_n]$ at birth draws a productivity parameter from a common distribution and learns its type, that is, its marginal cost a_i. The production and the export decision for each firm in country n occur after observing a_i. If a firm decides to produce, it has to pay the fixed cost of setting up production, denoted by f; if it decides to export, it has to pay the fixed cost of exporting, denoted by κ_{mni}. The price of the input bundle for country n is ω_n. Technology is linear in the input bundle: to produce one unit of output, a firm with a marginal cost a_i requires a_i units of the input bundle, i.e., if there is only one input, labour, then the cost of the input bundle is the wage $\omega_n = w_n$. Alternatively, in an economy with both labour and capital and a Cobb-Douglas production function, the cost of the input bundle is $\omega_n = w_n^{\alpha} r_n^{1-\alpha}$, where r_n is the return to capital in country n. Regardless of its productivity, each firm faces a residual demand curve with constant elasticity ε given by equation (1) and thus chooses the same profit maximizing markup equal to $\frac{\varepsilon}{\varepsilon-1}$, and sets its price as:

$$p_{ni} = \frac{\varepsilon}{\varepsilon - 1}\omega_n a_i$$

and supplies a quantity equal to,

$$\frac{Y_n}{P_n^{1-\varepsilon}} \left[\frac{\varepsilon}{\varepsilon - 1}\omega_n a_i\right]^{\varepsilon}$$

and earns revenues from domestic sales, D_i:

$$D_i = \frac{Y_n}{P_n^{1-\varepsilon}} \left(\frac{\varepsilon}{\varepsilon - 1}\omega_n a_i\right)^{1-\varepsilon}$$

which is obtained by substituting $p_{ni} = \frac{\varepsilon}{\varepsilon-1}\omega_n a_i$ into equation (1) and recognizing that revenue $D_i = p_{ni} c_{ni}$. (Note that firm's revenues fall with a_i.)

Decision to supply

We define $D_i = M_n \times B_i$ where $M_n \equiv \frac{Y_n}{P_n^{1-\varepsilon}} \left(\frac{\varepsilon}{\varepsilon-1}\omega_n\right)^{1-\varepsilon}$ is a measure of the size of the domestic demand, and $B_i \equiv a_i^{1-\varepsilon}$ is the firm-specific productivity term. The total variable profits are a constant multiple of D_i:

$$\pi_D^V(a_i) = \frac{D_i}{\varepsilon}$$

The zero cutoff profit condition requires that, given firm i's draw of productivity, a_i, the firm will only choose to supply if its variable profits cover the fixed cost of setting up production, that is if $\pi_D^V(a_i) \geq f$. Firms drawing a productivity level allowing them to make nonnegative profits will supply; otherwise, they will not supply. (This cutoff also determines the set of goods supplied to the market.) Now, we can define the minimum firm size, which is observed in this economy as: $\underline{D} = \varepsilon f$ and the marginal cost above which the firm will not operate:

$$a_{nn} = \left(\frac{M_n}{\varepsilon f}\right)^{\frac{1}{\varepsilon-1}}$$

Exporting from country n to country m involves additional fixed and variable costs[5]: to export from country n to country m, a firm i has to pay the fixed cost of exporting, κ_{mni} which is firm-variant, and a per-unit iceberg transport cost of exporting τ_{mn}>1. An iceberg transport cost of exporting greater to unity indicates that τ units of a good must be shipped in order

[5] [18] points out the evidence that the firms incur not only in per-unit costs such as transport costs and tariffs, but also in significant fixed costs associated with market research, adaptation of the product to the foreign standards, setting up distribution channels, etc.

for 1 unit to arrive at destination. The consumer in country m also maximizes

$$max \left[\int_{J_m} c_{mi}^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

s.t.

$$\int_{J_m} p_{mi} c_{mi} di = Y_m,$$

which is identically symmetric to the consumer in country n. Given the demand in country m $c_{mi} = \frac{Y_m}{P_m^{1-\varepsilon}} p_{mi}^{-\varepsilon}$ and the marginal cost equal to $\tau_{mn}\omega_n a_i$, the firm i in country n charges the profit maximizing price $p_{mi} = \frac{\varepsilon}{\varepsilon-1}\tau_{mn}\omega_n a_i$ and earns export revenues X_i given by:

$$X_i = \frac{Y_m}{P_m^{1-\varepsilon}} \left[\frac{\varepsilon}{1-\varepsilon} \tau_{mn}\omega_n a_i \right]^{1-\varepsilon}$$

m's market size that is faced by the firm i in country n is:

$$M_m^* \equiv \frac{Y_m}{P_m^{1-\varepsilon}} \left[\frac{\varepsilon}{1-\varepsilon} \tau_{mn}\omega_n \right]^{1-\varepsilon}$$

and the variable profit from exporting is given by $\frac{M_m^* B_i}{\varepsilon}$ and the exporting decision (the export cutoff) is thus based on:

$$\frac{M_m^* B_i}{\varepsilon} \geq \kappa_{mni} \tag{2}$$

The fixed cost of exporting is stochastic and varies for each firm, so each firm will have different export profit values, above which they decide to export. The free entry condition and the zero cutoff profit condition identify a unique average productivity level and an average profit level. The equilibrium average productivity level determines in turn, the export productivity cutoff as well as the average productivity levels and the ex ante successful entry and export probabilities. The free entry condition and the aggregate stability condition ensure that the aggregate payment to the labour used for investment equals the aggregate profit level. Thus, the aggregate revenue remains exogenously fixed by the size of the labour force. The average firm revenue is determined by the zero cutoff profit and free entry conditions, which pins down the equilibrium mass of incumbent firms, which determines in turn, the mass of variety available in each country. Here, we do not aim to solve the model; we will only derive the analytical power law relationship described in [11] using this model.

2.1 Power Law in Firm Size in Open and Closed Economy

In this section, we will first demonstrate that firm sales S_i follow a power law in a closed economy, where the firm productivity is Pareto and then describe the two mechanisms (which are analyzed in [11]) by which exporting flattens the distribution and makes it more fat-tailed in the presence of international trade.

Firm sales, S_i follow a power law, if the number of firms with a size S_i greater than a value s is proportional to $\frac{1}{s^\alpha}$ for some positive α, that is:

$$Pr(S_i > s) = Cs^{-\alpha},\ \alpha > 0. \tag{3}$$

As in [11], we assume that B_i follows a Pareto distribution with exponent

α, which amounts to the same thing as assuming that firm productivity is Pareto in this setup. Let us assume that firm productivity is given by $\frac{1}{a} \sim$ Pareto(b, θ), and its cumulative distribution function by: $Pr\left(\frac{1}{a}<y\right) = 1 - \frac{b}{y}^{\theta}$. In autarky, it can be written:

$$
\begin{aligned}
Pr(S_i>s) &= Pr(D_i>s) \\
&= Pr(M_n B_i>s) \\
&= Pr\left(a_i^{1-\varepsilon}>\frac{s}{M_n}\right) \\
&= Pr\left(\left(\frac{1}{a}\right)^{\varepsilon-1}>\frac{s}{M_n}\right) \\
&= Pr\left(\frac{1}{a}>\left(\frac{s}{M_n}\right)^{\frac{1}{\varepsilon-1}}\right) \\
Pr(S_i>s) &= (b^{\varepsilon-1}M_n)^{\frac{\theta}{\varepsilon-1}} s^{\frac{\theta}{\varepsilon-1}} \qquad (4)
\end{aligned}
$$

Here, $C = (b^{\varepsilon-1}M_n)^{\frac{\theta}{\varepsilon-1}}$ and $\alpha = \frac{\theta}{\varepsilon-1}$, so we can write $S_i =$Pareto$\left(b^{\varepsilon-1}M_n, \frac{\theta}{\varepsilon-1}\right)$.

Our next step is to analyse the power law relationship in presence of international trade. There are two mechanisms, proposed by [11], which allow to explain the flattening of the distribution due to the exporting behaviour. The first mechanism is based on the hierarchy of firm productivities: the more productive the firm, greater the number of foreign markets it serves (entry into progressively more foreign markets). The second mechanism proposed by [11] is based on stochastic export costs that vary by firm, allowing them to obtain a number of analytical results about the distribution of firm sizes, which is affected systematically by exporting behaviour.

Mechanism 1. For simplicity, we assume that the fixed cost of exporting is same for each firm i, i.e. $\kappa_{mni} = \kappa_{mn}$, $\forall i$. From equation (2), we can deduce that, each firm will have different number of foreign markets to serve, given its productivity draw. A firm in market n will first serve the markets with lowest τ_{mn} and κ_{mn} (that is the closest markets), high Y_m (markets with larger size), and high $P_m^{1-\varepsilon}$ (high prices/ lower competition). When we order the firms according to their productivity (marginal cost), we will see that the firms with the highest productivity serve m number of markets (if we assume m to be the number of world markets in total), and the firms with the next highest productivity level will serve $m-1$ number of markets and the firms with the lowest productivity will serve the domestic market and so on.

An ordering of firms in this fashion (see Figure 3) implies different power law exponents for exporting firms than that of non-exporting firms: the slope of the power law is constant and equal to $\frac{\theta}{\varepsilon-1}$ for non-exporting firms (equal to the autarky exponent). But for exporting firms, the CDF is systematically and parallely shifted upwards, due to the additional export markets they serve. That is, the more productive the firm, greater the number of markets it serves, and more fat-tailed the distribution becomes (see Figure 2).

Mechanism 2. For conveniency, we assume there is only one export market m. We define ϕ as the ratio of the foreign market size relative to the domestic one:

$$\phi \equiv \frac{M_m^*}{M_n} = \tau_{mn}^{1-\varepsilon} \frac{Y_m}{Y_n} \left(\frac{P_n}{P_m} \right)^{1-\varepsilon}$$

The exporting decision (2) can be rewritten as a function of domestic sales (and omitting the subscript mn from κ_{mni}):

$$\frac{\phi D_i}{\varepsilon} \geq \kappa_i$$

We define an export probability function $H(x)$: a firm with domestic sales D_i exports with probability $H(D_i)$, which is weakly increasing with D_i.

$$H(x) = Pr\left(\kappa_i \leq \frac{\phi x}{\varepsilon}\right)$$

Export sales X_i can now be rewritten as:

$$X_i = M_m^* B_i = \phi D_i$$

and

$$X_i = \begin{cases} 0 & \text{if } \frac{\phi D_i}{\varepsilon} < \kappa_i; \text{probability } 1 - H(D_i) \\ \phi D_i & \text{if } \frac{\phi D_i}{\varepsilon} \geq \kappa_i; \text{probability } H(D_i). \end{cases} \tag{5}$$

Recalling that $S_i = D_i + X_i$, total sales S_i can be written as:

$$S_i = \begin{cases} D_i & \text{if } \frac{\phi D_i}{\varepsilon} < \kappa_i; \text{probability } 1 - H(D_i) \\ (1+\phi) D_i & \text{if } \frac{\phi D_i}{\varepsilon} \geq \kappa_i; \text{probability } H(D_i). \end{cases} \tag{6}$$

Proposition 2.1 *The densities of domestic sales D_i, export sales X_i and total sales S_i are given by:*

$$p_D(x) = kx^{-\alpha-1} 1_{x > \underline{D}}$$

$$p_X(x) = Kx^{-\alpha-1} H\left(\frac{x}{\phi}\right) 1_{x > \phi \underline{D}} \tag{7}$$

$$p_S(x) = kx^{-\alpha-1}\left[1 - H(x) + H\left(\frac{x}{1+\phi}\right)(1+\phi)^{\alpha}\right] 1_{x > (1+\phi)\underline{D}} + kx^{-\alpha-1} 1_{\underline{D} < x < (1+\phi)\underline{D}} \tag{8}$$

where $k = \alpha \underline{D}^{\alpha}$, K is a constant ensuring $\int p_X(x)dx = 1$, and $1_{\{.\}}$ is the *indicator function which is equal to* 1 *if the statement in the brackets is true and else zero.*

This proposition implies that the distributions of total sales and export sales are systematically different due to the presence of exporting behaviour, under the assumption that the productivity distribution is Pareto. So, fitting a simple power law equation to the total sales data, whose theoretical distribution is described in (8) will be misleading. The common practice of fitting a power law on total sales will give incorrect estimates for α, which is in turn used to calibrate the parameter combination, $\frac{\theta}{\varepsilon-1}$.

Proof: From (5), the probability of exports conditioned on domestic sales D_i is:

$$P(X_i{>}0|D_i) = H(D_i)$$

We assume that the distribution of baseline sizes is:

$$p_D(x) = kx^{-\alpha-1}1_{x>\underline{D}} \tag{9}$$

where k is an integration constant, $k = \alpha\underline{D}^{\alpha}$.

We next calculate the distribution of exports. We calculate the expected value of $g(X)$, an arbitrary function. First, given (5),

$$E[g(X_i)|D_i] = (1 - H(D_i))g(0) + H(D_i)g(\phi D_i)$$

Hence,

$$\begin{aligned} E[g(X_i)] &= E[E[g(X_i)|D_i]] \\ &= E[(1-H(D_i))g(0)+H(D_i)g(\phi D_i)] \\ &= \int_D (1-H(D))g(0)p_D(D)dD + \int_D H(D)g(\phi D)p_D(D)dD \end{aligned}$$

$$E[g(X_i)] = \left(\int_D (1-H(D))p_D(D)dD\right) g(0) + \int_{x>0} \left(H\left(\frac{x}{\phi}\right) p_D\left(\frac{x}{\phi}\right)\frac{1}{\phi}\right) g(x)dx \quad (10)$$

Equation (10) implies that x has a point mass $\int_D (1-H(D))p_D(D)dD$ on $X=0$, and a density $H\left(\frac{x}{\phi}\right) p_D\left(\frac{x}{\phi}\right)\frac{1}{\phi}$ for $x>0$. Hence, the density associated with the restriction of the exports to $X>0$ is

$$p_X(x) = k' p_D\left(\frac{x}{\phi}\right) H\left(\frac{x}{\phi}\right)\frac{1}{\phi}$$

for a constant k' such that $\int_{x>0} p_X(x)dx = 1$. Given Equation (9),

$$p_X(x) = Kx^{-\alpha-1} H\left(\frac{x}{\phi}\right) 1_{x>\phi\underline{D}}$$

for a constant $K = k'\phi^\alpha k$. We can calculate the distribution of S_i using a similar approach. From Equation (6), with same reasoning as for exports gives:

$$p_S(x) = p_D(x)(1-H(x)) + p_D\left(\frac{x}{1+\phi}\right) H\left(\frac{x}{1+\phi}\right)\frac{1}{1+\phi}$$

With the Pareto specification for D:

$$p_S(x) = kx^{-\alpha-1}\left[1 - H(x) + H\left(\frac{x}{1+\phi}\right)(1+\phi)^{\alpha}\right]1_{x>(1+\phi)\underline{D}} + kx^{-\alpha-1}1_{\underline{D}<x<(1+\phi)\underline{D}} \quad (11)$$

We see a Pareto shape in the tails, but modulated by the export probability function H. $\blacksquare$

We will now present an auxiliary prediction. Suppose that the distribution of fixed exporting cost κ_i is Pareto, with an upper cut: $H\left(\frac{x}{\phi}\right) = k''x^{\beta}$, for $x<x^*$ and some k'', and $H\left(\frac{x}{\phi}\right) = k''(x^*)^{\beta}$ for $x>x^*$. Then, from Equation (7), we deduce that the distribution of export sales is given by:

$$p_X(x) \propto \begin{cases} x^{-\alpha-1+\beta} & \text{for } x<x^* \\ x^{-\alpha-1} & \text{for } x \geq x^*. \end{cases}$$

and the scaling parameter of X is given by:

$$\alpha_X(x) = \begin{cases} \alpha - \beta & \text{for } x<x^* \\ \alpha & \text{for } x \geq x^*. \end{cases}$$

When H has a high slope, the power law exponent of X is lower than that of domestic sales, i.e. when the fixed cost of exporting is high (the selection effect will be stronger), there will be fewer small exports; and when the H function saturates, the power law exponent of exports converge to α, that is the exponent for domestic sales (see Figure 4).

In this section, we have explained the two mechanisms by which the scaling parameter for exporting firms is systematically lower (in absolute value) than that for non-exporting firms. In the next sections, we will test whether the data is in accord with the theoretical findings.

3 Estimating and Testing a Power Law Distribution

In this section, we will discuss how a power law distribution should be fitted to the empirical data in an accurate way and how to test whether a power law hypothesis is plausible for the data in question. Many studies in the literature use different estimation and hypothesis testing methods, and in some cases the plausibility of the power law hypothesis is not formally tested. The method most employed in determining the power law exponent is graphical analysis of the log of the ranked data or a frequency histogram of the data, followed by a least-squares linear regression. However, it is argued by many studies that OLS estimates are subject to systematic and potentially large errors and some other serious problems; and that maximum likelihood estimation produces more accurate and unbiased estimates than does the OLS method [7].

3.1 Linear regression and power law

In the literature, the scaling parameter is usually inferred by fitting a straight line fit to: (i) the slope of a log-transformed histogram, (ii) the slope of a histogram with logarithmic bins (increasing bin widths), (iii) the slope of the CDF calculated with constant width bins, and (iv) the slope of the CDF calculated without any bins. Usually, the R^2, the fraction of explained variance, is reported as an indicator of the quality of the fit; and the standard error of the slope estimate as a measure of the uncertainty of the scaling parameter. These approaches are unreliable as the assumptions underlying the least-squares regression do not hold.

An OLS estimation based on the plot of the empirical probability distribution $\hat{p}(x)$ on a double-logarithmic scale systematically underestimates the power law exponent because of the sparseness of data in the tail of the distribution. To deal with the lack of data points in the tail of the empirical distribution, two methods are employed in the literature: logarithmic binning and estimating the empirical cumulative distribution $\hat{P}(x)$, instead of $\hat{p}(x)$: (i) Logarithmic binning reduces the noise in the tail of the empirical distributions $\hat{p}(x)$ and $\hat{P}(x)$ by merging data points into groups. As a consequence of binning, the width of the distribution of the estimate is reduced. (An important drawback of this method is that binning throws away information.) (ii) Estimating the CDF, which is less sensitive to the noise in the tail of the distribution, gives much better estimates for the exponent.[6] However, all graphical methods have a common serious flaw: When OLS estimates are derived, it is assumed that the standard deviation of the distribution of the error in y_i is the same for all data points (x_i, y_i) [7]. Here, this is not the case, which causes the following limitations:

All graphical methods for estimating the scaling parameter are based on a linear least squares fit of some empirical data points $(x_1, y_1), (x_2, y_2), ..., (x_M, y_M)$

[6]It should be noted that this method is inefficient, because the values are not independent and the inferential assumptions for the OLS regression fail.

to the equation

$$y(x) = a_o + a_1 x.$$

The linear least squares fit minimizes the residual

$$\sum_{i=1}^{M} \hat{u}^2 = \sum_{i=1}^{M} (y_i - a_0 - a_1 x_i)^2$$

Estimates $\hat{a_0}$ and $\hat{a_1}$ for the parameters a_0 and a_1 are given by:

$$\hat{a_1} = \frac{\sum_{i=1}^{M} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{M} (x_i - \bar{x})^2}$$

and

$$\hat{a_0} = \bar{y} - \hat{a_1}\bar{x}$$

Derivation of the parameter estimates are based on several assumptions regarding the data points (x_i, y_i): it is assumed that there are no statistical uncertainties in x_i, but y_i may contain some statistical error. The errors in different y_i are independent and identically distributed random variables with zero mean and the standard deviation of the errors is independent of x_i [23]. However, these conditions are not fulfilled for various graphical methods for estimating the scaling parameter (i.e. the empirical values are highly correlated and the statistical variation in the values is not constant, but increases rapidly with the degree, which is caused by the logarithmic nature of the plot and the decreasing probabilities [17].)

1. Calculation of standard errors

The exact normality of the OLS estimators hinges crucially on the normality of the distribution of the error in the population. If the errors are random draws from some distribution other than the Gaussian, the OLS estimates will not be normally distributed, which means that the t statistics will not have t distributions and the F statistics will not have F distributions. This is a potentially serious problem because our inference hinges on being able to obtain p-values from the t or F distributions. When we perform a linear regression on the logarithm of a histogram, the error will not be Gaussian (though independent): The error in the frequency estimates $p(x)$ themselves is Gaussian, but the error in their logarithms is not: for $\ln p(x)$ to have Gaussian fluctuations, $p(x)$ would have to have log-normal fluctuations, which would violate the central limit theorem[7]. Thus, the ordinary formula for the calculation of the standard error of the estimate is in this case incorrect and not applicable.

When we consider the case of fitting a straight line to a CDF, the error in the individual values $P(x)$ is Gaussian (since it is the sum of independent Gaussian variables), but the error in the logarithm is not. Moreover, the independency assumption is violated, since the statistical errors in y_i are not independent any more (i.e. the adjacent values of the CDF are highly correlated, because $P(x) = P(x+1) + p(x)$)[4]. CDF fits are empirically more accurate for determining α, however this is not due to the fact that the assumptions of the linear regression are valid, but because the statistical fluctuations in the CDF are much smaller compared to those in the PDF. The standard error on the scaling parameter becomes smaller, but this does not mean that the estimate of the error is better; it is smaller because of the failure to account for correlations [4].

2. **Validation** Using R^2 as an indicator of the quality of fit may in most cases be misleading. We may observe many different distributions, which

seem to follow a power law over some of their range, such as the log-normal or the streched exponential, and which may both give quite high R^2 values, but are indeed not power law distributions. Even when the fitted distribution approximates a power law quite poorly, we may still observe high R^2 values, since the fit accounts for a significant fraction of the variance. Although, we should note that a small R^2 may be informative and may allow us reject the power-law hypothesis. But, in most of the cases, a low R^2 will rarely be encountered, therefore it will tell us nearly nothing about the goodness-of-fit of a power law[7].

3. **Regression lines are not valid distributions** The PDF must integrate to 1 over its defined range, however this constraint is not taken into account by the OLS regression, thus the regression line will not respect this constraint[7].

3.2 Maximum likelihood estimation

Maximum likelihood method is a more reliable approach than the graphical methods mentioned above. Suppose that $f(x;\alpha)$ denote the probability density function of a random variable x, conditioned on a single parameter α. The joint density of n independent and identically distributed observations from this process is the product of the individual densities:

$$p(x_1, ..., x_n; \alpha) = \prod_{i=1}^{n} p(x_i; \alpha) = L(\alpha; x_1, ..., x_n)$$

The joint density is the likelihood function, defined as a function of the unknown parameter, α, where $x_1, ..., x_n$ is used to indicate the collection of sample data. For mathematical convenience, the logarithm of the likelihood function is maximized, as its maximum is in the same place:

$$\ln L(\alpha; x_1, ..., x_n) = \sum_{i=1}^{n} \ln p(x_i; \alpha)$$

The maximum likelihood estimator of α is that value of α that maximizes the sample likelihood function, L. We maximize the log-likelihood function by differentiating it with respect to the unknown and equating the resulting derivative to zero. The resulting value of the estimator is called the maximum likelihood estimator. Under some regularity conditions, the maximum likelihood estimator has the following asymptotic properties[13]:

1. Consistency: $plim \quad \hat{\alpha} = \alpha$ (where α is the true unknown parameter)

2. Asymptotic normality: The estimator $\hat{\alpha}$ is asymptotically normal with mean α and variance $\frac{(\alpha-1)^{-2}}{n}$.

3. Asymptotic efficiency (no consistent alternative with less variance): $\hat{\alpha}$ is asymptotically efficient and achieves the Cramer-Rao lower bound for consistent estimators.

4. Invariance: If $\hat{\alpha}$ is the maximum likelihood estimator for α, and if $g(\alpha)$ is any transformation of α, then the maximum likelihood estimator for $g(\alpha)$ is $g(\hat{\alpha})$.

Maximum likelihood estimators for the power law

We derive the maximum likelihood estimator for the power law distribution $p(x) = \frac{\alpha-1}{x_{min}} \left(\frac{x}{x_{min}}\right)^{-\alpha}$, where α is the scaling parameter and x_{min} indicates the threshold value above which the distribution obeys a power law. We will calculate the unknown parameter α in such a way that the probability of observing the sample values is maximized. Given $p(x)$ and assuming that the observations are independent and identically distributed, we can write the likelihood function for $x_i \geq x_{min}$:

$$p(x;\alpha) = \prod_{i=1}^{n} \frac{\alpha - 1}{x_{min}} \left(\frac{x_i}{x_{min}} \right)^{-\alpha}$$

This probability is called the likelihood of the data given the model. The convention is to work with the logarithm of the likelihood function:

$$\begin{aligned} \mathcal{L} &= \ln p(x;\alpha) = \ln \prod_{i=1}^{n} \frac{\alpha - 1}{x_{min}} \left(\frac{x_i}{x_{min}} \right)^{-\alpha} \\ &= \sum_{i=1}^{n} \left[\ln(\alpha - 1) - \ln x_{min} - \alpha \ln \frac{x_i}{x_{min}} \right] \\ &= n \ln(\alpha - 1) - n \ln x_{min} - \alpha \sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}. \end{aligned} \tag{12}$$

Differentiating $\mathcal{L}$ with respect to α and equating to zero, we obtain the maximum likelihood estimator for α:

$$\hat{\alpha} = 1 + n \left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}} \right]^{-1}. \tag{13}$$

The standard error on $\hat{\alpha}$, which is derived from the width of the likelihood maximum is

$$\sigma = \frac{\hat{\alpha} - 1}{\sqrt{n}} + O(1/n) \tag{14}$$

where the higher-order correction is positive[7].

Estimating the lower bound on power law behaviour

The power law is rarely encountered over the entire range of real-world phenomena and usually is not a good description for smaller values of the variable of interest. We can easily see that $p(x) = Cx^{-\alpha}$ diverges when x tends to zero, so we expect that a distribution follows a Power law only in its upper tail, after a certain minimum threshold x_{min}. Beyond this minimum threshold, the distribution deviates from the Power law. Thus, to obtain an accurate estimate of the scaling parameter, we need to ignore any observation below this lower bound. If we choose too low an x_{min}, then the estimate of the scaling parameter will be biased, since we are trying to fit a power law model to a non-power-law data. Again, if we choose too high an x_{min}, then we will be losing observations, that actually fall into the power law branch of the distribution, thus increasing the statistical error on the scaling parameter.

In the literature, the most employed methods for choosing x_{min} are: (i) to infer visually the point beyond which the PDF or CDF of the empirical distribution becomes straight on a log-log plot, (ii) to plot $\hat{\alpha}$ as a function of x_{min} and identify the point beyond which the value appears stable.

The method proposed by [7] is as follows: we choose the $\hat{x}_{min}$ that makes the probability distributions of the measured data and the best-fit power law model as similar as possible above $\hat{x}_{min}$. For quantifying the distance between two probability distributions, we use the Kolmogorov-Smirnov (KS) statistic, which is the maximum distance between the CDFs of the data and the fitted model:

$$D = max_{x \geq x_{min}} |S(x) - P(x)|$$

where $S(x)$ is the CDF of the data for the observations with value at least x_{min}, and $P(x)$ is the CDF for the power-law model that fits the data best above x_{min}. Then, $\hat{x}_{min}$ is the value of x_{min} that minimizes D. [7] find that the KS method gives good estimates for x_{min}.[7]

Testing the power law hypothesis

In the literature, the plausibility of the power law hypothesis is often tested in a qualitative way. If the data seems to be roughly straight on a log-log plot, or if a high R^2 is obtained, then one asserts that the data obey power law. However, this method may be misleading, as it may allow non-power-law distributions to be tagged as power-law distribution. In order to test the plausibility of the power law hypothesis, we use the method described in [7]: we sample synthetic data sets from a true power-law distribution, measure how far they fluctuate from the power-law form, and compare the results with similar measurements on the empirical data. If the empirical data set is much farther from the power-law form than the synthetic ones, then the power law is not a plausible fit to the data. This method employs the KS statistic as a measure of the distance between the distributions[8]. The intuition behind this method is that we need to distinguish the deviations due to the random nature of the sampling process from those that occur due to the non-power-law nature of the data.

[7] This method works also better than the the Bayesian information criterion method, where the maximum of the BIC with respect to the x_{min} gives the estimated value $\hat{x}_{min}$[7].

[8] It should be noted that, a non-power-law process may happen to generate a data set that is close to a power law, in which case this test will fail. However, the probability of this happening decreases with the sample size. See [7].

More precisely, we wish to know whether the power law hypothesis is plausible, given the data at hand. [7] use a goodness-of-fit test, which generates a p-value that quantifies the plausibility of the hypothesis. The distance between the distribution of the empirical data and the hypothesized model is compared with the distance measurements between the synthetic data sets and their own power-law fits. Then, the p-value is computed as the fraction of the synthetic distances that are larger than the empirical distance. In this case, if the p-value is large (close to 1), then we may say that the difference between the empirical data and the power-law model are solely due to statistical fluctuations; if the p-value is small, then we say that the power-law model is not a plausible fit for the observed data.

The procedure is as follows: first, a power-law model is fitted to the empirical data using the methods for estimating the lower bound and the scaling parameter explained above, then the KS statistics is calculated. Second, a large number of synthetic data sets are generated, which are drawn from a power law distribution with the exact same scaling parameter and lower bound as the distribution that best fits the observed data. Then, the KS statistic for each synthetic data set and its individual power-law fit is calculated. Finally, the p-value is calculated as the fraction of the time the resulting statistic is larger than that for the empirical data.

The generation of the synthetic data set in [7] is as follows: Suppose that the observed data set has n_{tail} observations for $x \geq x_{min}$ and n observations in total. With probability n_{tail}/n, a random number x_i is drawn from a power law with a scaling parameter $\hat{\alpha}$ and $x \geq x_{min}$; with probability $(n - n_{tail})/n$, they select one element uniformly at random from among the elements of the observed data set that have $x \leq x_{min}$ and set x_i equal to that element. Repeating the process for all $i = 1, \ldots, n$ they generate a complete synthetic data set that follows a power law above x_{min} but has the same non-power-law distribution as the observed data below x_{min}.

How many synthetic data sets to generate depends on how much accurate we want our p-values to be. If we wish our p-values to be accurate within about ϵ of our p-value, then we should generate at least $\frac{1}{4}\epsilon^{-2}$ synthetic data sets [7]. Here, we have generated 1000 such synthetic data sets, so in our case $\epsilon = 0.0158$.

Finally, following [7], we rule out the power law hypothesis if $p \leq 0.1$. However, it should be noted that, even if we obtain a *high* p-value, we cannot tell with absolute certainty that the power law is the correct distribution for the data. First because, there may well be other distributions that fit the data better than the power law, and second, when we have small n, it is harder to rule out the power law hypothesis; so a high p-value may be due to small n, in which case it should be treated with caution. In order to compare with alternative distributions, [7] propose using Vuong's Test to make a direct comparison between models and decide which one fits better. If this gives inconclusive results, then we should look at theoretical factors (take into consideration the industry dynamics for example) to make a sensible judgment about which distributional form is more reasonable.

4 Data and Empirical Methodology

4.1 Data description

The analysis is carried out with a large dataset of Turkish firms for 2009. The dataset is constructed using domestic sales data extracted from the Annual Industry and Service Database and export sales data from the Foreign Trade Database, both collected and provided by Turkstat for the year 2009. The agricultural sector is excluded from the dataset. For firms having more than 20 employees, full enumeration; for firms having less than 20 employees, sampling method is used. In total, the dataset includes 99.921 firms, which constitutes about 4% of the total number (2.483.300) of active firms. The number of exporting-firms is 14.231. The summary statistics are given in Table 13.

The power law model is usually not a good description for the data below some minimum threshold. So, we need to find the minimum threshold above which the power law model fits well to the data. In the literature, the most common method for this is just to look at the empirical cumulative distribution (or the PDF) by eye, and decide a minimum cutoff after which the plot becomes more or less linear. Another method is to plot the estimated α as a function of the minimum threshold and locate the point beyond which the estimated α appears relatively stable. Here, we use the method proposed by [7] to estimate the minimum threshold: we choose the value of $\hat{x}_{min}$ that makes the probability distribution of the empirical data and the best-fit power law model as similar as possible above $\hat{x}_{min}$. The measure used in order to quantify the distance between the CDFs of the data and the fitted model is the KS statistic: $D = \max_{x \geq x_{min}} |\hat{P}(x) - P(x; \hat{alpha}, \hat{x}_{min})|$ So the threshold estimate corresponds to the value of x_{min} that minimizes D:

$$\hat{x}_{min} = argmin_{x_{min}} \max_{x \geq x_{min}} |\hat{P}(x) - P(x; \hat{\alpha}, \hat{x}_{min})|$$

However, this method may be criticized for its somewhat circular reasoning: we use the goodness-of-fit of a power law model to identify the threshold which maximizes the ability of the power law to describe the data. Be that as it may, we use this method to estimate the minimum threshold, as we do not have a better option. The resulting minimum threshold may not be relevant economically, so we will not load it any economic meaning. This threshold is approximately TL[9] 30.000.000, it may be less or more depending on the considered dataset.

We estimate the power law using different samples: (i) First, we estimate the scaling parameter for all firms, without making a distinction between exporters and non-exporters and without omiting the non-tradeable sectors. (ii) Second, we take into account the exporting behaviour, for the whole sample. (iii) We estimate the scaling parameter using domestic sales only, for all firms. (iv) We repeat the same steps for only manufacturing firms. (v) We omit the non-tradeable sectors and estimate the power law for the remaining sample. (vi) We estimate the scaling parameter for exporters and non-exporters at sector-level.

[9]Turkish Lira

4.2 Empirical Methodology

Here, we use three different methods to estimate the scaling parameter. In addition to MLE, we also use the standard methods that are common in the literature in order to compare our findings with other studies. (i) First, we use the maximum likelihood method. We estimate the parameters x_{min} and α of the power law model using the method explained in section 3.2. Second, we calculate the goodness-of-fit between the data and the power law by generating many synthetic data sets from a true power-law distribution and measuring how far they fluctuate from the power-law form and finally comparing the results with similar measurements on the empirical data. Based on the resulting p-value, we decide whether we may or may not reject the plausibility of the power law hypothesis.

(ii) The second method is based on [3]. We regress the natural log of the complementary cumulative distribution function on the natural log of sales:

$$\ln Pr(S_i > s) = \ln(C) - \alpha \ln(s) \tag{15}$$

(iii) As the third method, we use the log-rank-log-size estimator proposed by [10], where they modify the regression with a shift of 1/2 (which they find to be as the optimal shift) in order to reduce the bias related to the OLS procedure. Accordingly, we rank firm sizes (sales) from largest (rank 1) to smallest (rank n), and denote their sizes as $S_{(i)}$. Then, we regress

$$\ln(i - 0.5) = \ln(C) + \hat{\alpha} \ln(S_{(i)}) + u_i \tag{16}$$

The standard error of the scaling parameter is given by $\sqrt{\frac{2}{n}}\hat{\alpha}$[10].

5 Results

In order to compare our findings with other empirical studies in the literature, we first do not consider exporting behaviour and do the analysis for all firms, using the three methods explained above (MLE, CDF, log-rank-log-size) to estimate the scaling parameter. The estimates are reported in Table 6. We see that our findings differ only slightly from those of [3] and [11]: [3] finds an α of 0.994 for U.S. firms, and [11] finds 1.017 for French firms (both with the CDF method), whereas we find an α of 1.072. The MLE delivers an estimate of 1.0271 ± 0.0186. Our results for each three method are presented graphically in Figure 5.

In Table 6, we give the p-value for the power-law model, which gives a measure of how plausible the power law is as a fit to the data. Here, the p-value is 0.2140 and passes the threshold of 0.10, but it should be noted that it is still questionable whether the power-law fit is the best fit. In order to say that the power law is the best fit, one has to compare among alternative distributions (such as the log-normal or the streched exponential), for example using Vuong's Test described in [7]. When we look at the plots taken respectively from [3] and [11] in Figure 12, we can see that there are some deviations in both tails, even though they are somewhat masked due to the binned nature of the data. However, in our plots (see Figure 5) the deviations from the power law in the upper tail are clearly visible, which raises some questions about the plausibility of a power-law model for firm size distribution. We will not go further on this issue for now; as we want to investigate whether the theoretical findings discussed in Section 2.1 are verified by the data.

The estimates of the scaling parameter for exporting and non-exporting firms are reported in Table 9. We easily see that, for each estimation method, the scaling parameter for exporting firms are lower (in absolute value) than that for non-exporting firms, i.e. the distribution of exporting firms is systematically more fat-tailed than the distribution of the non-exporting firms, which is consistent with the theoretical findings. We also find that the scaling parameter estimate for the overall sample (all firms) fall between the estimates for exporting and non-exporting firms, which is exactly as we expect. The results are presented graphically in Figure 7. The CDF of the exporting firms has a lower slope and is above the CDF of the non-exporting firms, which means that at each size cutoff, the number of the exporting firms is greater than the number of non-exporting firms (See Figure 6). As is already indicated in Section 2.1, we expect power law to be a bad fit for the exporting-firms sample, because of the deviations caused by the exporting behaviour. The p-value for exporting-firms sample is 0, so the power law model for exporters can be firmly ruled out. However, the p-value for non-exporting firms is 0.4880, and power law model seems satisfactory for non-exporters. Again, it should be noted that this does not mean that the power law is the best fit for the data and that one has to compare among the competing distributions.

The maximum likelihood estimates of scaling parameter in total sales, based on only manufacturing firms and only tradeable sectors are given respectively in Tables 7 and 8. When we take international trade into account, we obtain similar results. The scaling parameter for exporters is lower in absolute value than that for non-exporters, for both the manufacturing firms and the tradeable sample. Also, the p-value for exporters in tradeable sectors is also lower compared to the p-value for non-exporters.

As discussed in Section 2.1, we expect the scaling parameter for exporting firms to converge to the scaling parameter for non-exporting firms, as we increase the lower cutoff. This would be the case either because all the most productive (thus bigger) firms have entered all world markets or because the stochastic fixed exporting costs have an upper bound, above which all firms export. We have tested this finding by estimating the scaling parameter for exporters while moving the lower cutoff upwards. Figure 10 shows the scaling parameter estimates as a function of the lower cutoff. We see that the scaling parameter estimates increase with the lower sales cutoff, which is in line with the theoretical findings.

We have also estimated the power law in firm size for exporting and non-exporting firms at the industry level. Table 10 reports the relevant maximum likelihood estimates. Given the theoretical findings, we expect the results to exhibit the same pattern at the industry level, as at the aggregate level. Indeed, we observe that for each sector, the scaling parameter for exporters is smaller than that for non-exporters. Therefore, it is verified that exporting behaviour produces the same effects at industry-level, as at aggregate level.

Also in line with the theory, we expect to see greater deviations in the scaling parameter of sectors that are more open to trade. We test this prediction in the following way: We first estimate the power law in all sales and in domestic sales for each sector. Then, we calculate the percentage deviation for each sector, that is, the difference between the scaling parameter for domestic sales and for total sales of all firms for each sector. We then plot this against the overall sector openness, that is, the ratio of exports to total sales. As can be seen in Figure 11, the difference between the power law exponent for domestic sales and that for total sales increases with overall sector openness. To demonstrate the same prediction, we have done the same exercise considering the difference between the scaling parameter for non-exporting firms and that for all firms. Again, we plot this against the overall sector openness and find similar results: The percentage deviation in the scaling parameter increases with sector level openness. These results support the theoretical findings, according to which international trade systematically changes firm size distribution to make it more heavy-tailed.

6 Conclusion

In Section 2.1, we show that α, in autarky, the scaling parameter of the power law is equal to $\frac{\theta}{\varepsilon-1}$, where θ is the productivity distribution parameter (recall that $Pr\left(\frac{1}{a}<y\right) = 1 - \left(\frac{b}{y}\right)^{\theta}$) and ε is the elasticity of substitution between varieties. This has led researchers to use the estimates of α to determine $\frac{\theta}{\varepsilon-1}$. For example, [6] analyses the sensitivity of trade flows to trade barriers, using $\frac{\theta}{\varepsilon-1}$. He measures this as the regression coefficient of the log of rank on the log of sales (using Compustat data on US listed firms). [16] estimate the dispersion of firm size to capture the within-industry firm heterogeneity using total sales data. However, as is showed in Section 2.1, the presence of international trade alters the firm size distribution and the estimated α no

longer equals $\frac{\theta}{\varepsilon-1}$. Fitting a power law model on total sales, without taking international trade into account delivers too low estimates for α. In order to get correct estimates for $\frac{\theta}{\varepsilon-1}$, [11] propose two methods. The first one is to estimate α for non-exporting firms only, where theoretically α does correspond to $\frac{\theta}{\varepsilon-1}$. The second method they propose is to estimate the scaling parameter using the domestic sales for all firms, since the domestic sales obey power law with the $\frac{\theta}{\varepsilon-1}$ exponent.

Using the methods proposed in [11] and a large sample of Turkish firms, we have estimated the scaling parameter with three different estimation methods (MLE, CDF and log-rank-log-size). We have underlined the limitations of the OLS estimation when used for power law estimation and how MLE avoids these pitfalls. Following [7], we have also implemented a formal method to test the goodness-of-fit of a power law model for our data, and obtained a p-value of 0.2140 (for all firms, total sales) which exceeds the critical p-value 0.10. Therefore, we do not reject the power law hypothesis for firm size distibution. The p-values for the non-exporting and exporting firms are respectively 0.4880 and 0, which is consistent with the theory: Since exporting causes changes in the probability density, the simple power law describes the distribution of exporters slightly less well. However, the p-values exceeding 0.10 should be treated with caution. This does not mean that the power law model is the best description for the data in question and it still needs to be checked with Vuong's Test which competing distribution (for example, log-normal or streched exponential) is a better fit for the data [7]). Since the main objective of this paper is not whether the power law provides the best description or not, we did not go further on this issue.

The main results of the paper are the following: taking the presence of international trade into account, we have obtained estimates for the scaling parameter for exporters and non-exporters separately. We did this for the aggregate economy, for only manufacturing firms, for only tradeable sectors and at sector-level and systematically obtained lower scaling parameters for exporting firms, compared to non-exporting firms, for each different sample considered. This means that exporting behaviour produces the same effects at sector-level as at the aggregate level. We have also found a positive relationship between the deviations in the power law parameter (the difference between the power law exponent for domestic sales and that for total sales) and sector-level openness (the ratio of exports to total sales in the sector), which is consistent with the theory. All these findings provide support for the prediction that international trade systematically changes the firm size distribution to make it more fat-tailed and that ignoring the presence of international trade yields estimates that are too low.

References

[1] Arkolasis, C., *Market Penetration Costs and the New Consumers Margin in International Trade*, 2008, Mimeo, Yale University.

[2] Atiyas, I., Bakis, O., *Obstacles to growth in Turkey*, TUSIAD-T/2011/11/519 (in Turkish).

[3] Axtell, R., *Zipf Distribution of U.S. Firm Sizes*, Science, September 2001, 293 (5536), 1818-1820. IMF Working Papers 10/109, 2010. International Monetary Fund.

[4] Bauke, H., *Parameter Estimation for Power-Law Distributions By Maximum Likelihood Methods*, The European Physical Journal B, Volume 58, Number 2 (2007), 167-173.

[5] Champernowne, D., *A Model of Income Distribution*, 1953, Economic Journal, 83, 318-351.

[6] Chaney, T., *The Intensive and Extensive Margins of International Trade*, The American Economic Review, Vol. 98, No. 4 (Sep., 2008), pp. 1707-1721.

[7] Clauset, A., Shalizi, C.R., and Newman, M.E.J. *Power-law distributions in empirical data*, SIAM Review 51(4), 661-703 (2009).

[8] Eaton, J., Kortum, S., and Kramarz, F., *An Anatomy of International Trade: Evidence from French Firms*, Econometrica, Vol. 79, No. 5 (September, 2011), 1453-1498.

[9] Gabaix, X., *The Granular Origins of Aggregate Fluctuations*, August 2009, NBER Working Paper No. 15286.

[10] Gabaix, X., Ibragimov, R., *Rank−1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents*, 2008, Journal of Business and Economic Statistics.

[11] di Giovanni, J., Levchenko, A. and Ranciere, R., *Power Laws in Firm Size and Openness to Trade: Measurement and Implications*, IMF Working Papers 10/109, 2010. International Monetary Fund.

[12] di Giovanni, J., Levchenko, *International Trade and Aggregate FLuctuations in Granular Economies*, February 2009. RSIE Discussion Paper 585.

[13] Greene, W. H., *Econometric Analysis*, 2003, Prentice Hall.

[14] Hall, B. H., *The relationship between firm size and firm growth in the U.S. manufacturing sector*, Journal of Industrial Economics, 35, pp. 583-606, 1987.

[15] Hart, P. E., Prais, S. J., *The analysis of business concentration: a statistical approach*, Journal of the Royal Statistical Society, 119, pp. 150-191, 1956.

[16] Helpman, E., Melitz, M., Yeaple, S. R., *Export versus FDI with heterogeneous firms*, American Economic Association, Vol. 94, No. 1 (Mar., 2004), pp. 300-316.

[17] Jones, J. H., Handcock, M. S., *An Assessment of Preferential Attachment as a Mechanism for Human Sexual Network Formation*, The Royal Society, June 2003 vol. 270 no. 1520 1123-1128.

[18] Melitz, M., *The Impact of Trade on Intra-Industry Reallocations and*

Aggregate Industry Productivity, Econometrica, Vol. 71, No. 6 (Nov., 2003), pp. 1695-1725

[19] Rodrik, D., *Structural Change and Development*, TEPAV, December 2011 (in Turkish).

[20] Sutton, J., *Gibrat's Legacy*, Journal of Economic Literature 35, 40-59, 1997.

[21] Stanley M. H. R., Buldryev, S.V., Havlin, S., Mantegna, R., Salinger, M., Stanley, E., *Zipf plots and the size distribution of firms*, Economic Letters 49, 1995, 453-457.

[22] Quandt, R., *On the size distribution of firms*, American Economic Review 56, 416-432.

[23] Wooldridge, J. M., *Introductory Econometrics: A Modern Approach*, 2009, South-Western College Publications.

[24] Zipf, G., *Human Behavior and the Principle of Least Effort*, Cambridge, Mass: Addison-Wesley,1949.

A An Overview of the Turkish Industry

Turkish economy is characterized by the prevalence of SMEs, which account for approximately 99% of the total number of firms, 76.7% of total employment, 26.5% of investment and 10% of exports [2]. According to the Annual Industry and Service Statistics, $95,62\%$ of firms are micro ($1-9$ employees), $3,78\%$ small ($10-49$ employees), $0,50\%$ medium ($50-249$ employees) and $0,10\%$ are large-scale (250+ employees). An average SME has 3 employees, whereas the large-scaled businesses have 735. In Turkey, SMEs are characterized by their slow growth, low innovation levels, difficulties in accessing finance and new markets, and they lack skilled labour [2]. All these factors are counted among the obstacles to their upsizing.

The GDP and employment shares of the main sectors (as of 2011) are given in Table 1. The low share of industry in GDP is attributed to the failure of industrial policies implemented between $1950-1975$. The failure of these import substitution industrialization policies have led to an import dependent industry. Turkey has also suffered from low levels of funding for R&D activities due to lack of political support and lack of resources to support the proper development of science and technology policies. The evolution of the share of R&D expenditures in GDP is given in Table 4 (EU-27 average is also given for comparison). Despite the increasing trend, it is still quite low compared to EU-27. This problem manifests itself in the technology structure of exports, which is given in Table 3.

Table 2 presents the main indicators for Turkish manufacturing industry for the period 1998 and 2011. The share of manufacturing industry in GDP was around 23.55% on average for the period considered. We observe that

Table 1: GDP and Employment Shares of Main Sectors (%), 2011

	Share in GDP	Share in Employment
Agriculture	9.2	25.5
Industry	26.9	19.5
Services	63.9	55.1

Source: TURKSTAT

the share of the manufacturing industry in exports has increased, however, the low-tech sectors' share continued to remain dominant (see Table 3. This is mostly due to the low levels of R&D expenditures.

Atiyas et al. (2010)[2] argue that labour productivity and total factor productivity contribute little to growth in the Turkish economy. Rodrik [19] finds that the total factor productivity growth was negative between 1990 and 2003, while the labour productivity growth was positive and of 0.9%. Rodrik [19] reports the labour productivity growth between $1999 - 2005$ as 5%, and that the 40% of this growth is due to the allocation of labour force between sectors. As seen in Figure 1, the differential in productivity levels (value added per unit of labour) between agriculture and manufacturing industry is significant (it would make better sense to make a comparison between agriculture and manufacturing, since the utilities industry is capital-intense). As it is observed, the average productivity of manufacturing is almost three times greater than that of agriculture. With such huge productivity differentials, we may conclude that an allocation between sectors will result in nontrivial productivity gains.

Figure 1: Inter-industry productivity differentials, 2008 [19]

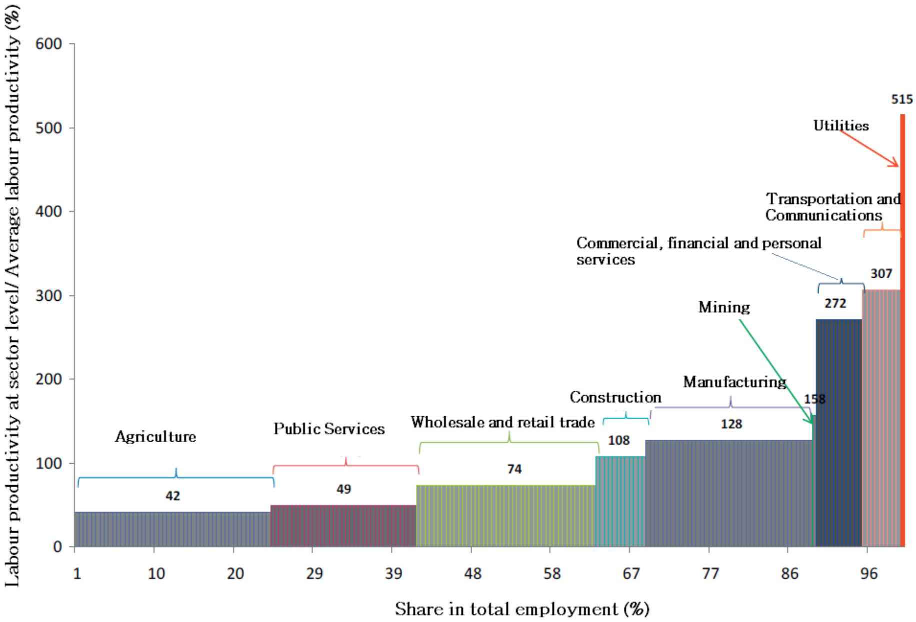


Table 2: Main Indicators for Manufacturing Industry

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Share in GDP	23.91	23.49	23.52	23.04	22.33	22.99	23.53	23.48	23.81	24.03	23.83	23.23	24.18	24.39
Share in total exports (% of total)	89.2	90.1	91.9	92	93.5	93.9	94.3	93.7	93.8	94.2	94.8	93.4	-	-
Capacity utilization rate	76.5	72.4	75.9	70.9	75.4	78.4	81.3	80.3	81	80.27	76.69	65.27	72.60	75.40
Exports/Imports (%)	–	–	–	–	–	–	–	70.7	70.4	72.7	78.8	83	83	–
Value added (% of GDP)	26	24	23	21	20	20	20	20	20	19	18	17	18	–

Source: TURKSTAT

Table 3: Technology Structure of Manufacturing Industry Production and Exports (%)

	Turkey						EU Exports
Technological intensity	Production			Exports			
	2000	2002	2005	2000	2002	2005	2003
High	$5,9$	$5,1$	$6,3$	$7,8$	$6,2$	$6,0$	$21,5$
Mid-High	$22,5$	$18,2$	$25,3$	$20,4$	$24,3$	$28,5$	$41,9$
Mid-Low	$30,4$	$26,7$	$27,0$	$20,5$	$22,8$	$26,9$	$15,9$
Low	$41,2$	$50,0$	$41,4$	$51,3$	$46,8$	$38,7$	$20,7$
Total	$100,0$	$100,0$	$100,0$	$100,0$	$100,0$	$100,0$	$100,0$

Source: State Planning Organization, Ninth Development Plan (2006)

Table 4: Total R&D Expenditure as % of GDP

	Turkey	EU-27
2000	0.48	1.86
2001	0.54	1.87
2002	0.53	1.88
2003	0.48	1.87
2004	0.52	1.83
2005	0.59	1.83
2006	0.60	1.85
2007	0.72	1.85
2008	0.73	1.92
2009	0.85	2.01
2010	0.84	2.00

Source: State Planning Organization, Ninth Development Plan (2006)

Table 5: Average TFP and Labour Productivity Growth, %

	TFP Growth	Labour Productivity Growth
$1960-1980$	1.75	3.2
$1980-1990$	1.45	2.5
$1990-2003$	-0.70	0.9

Source: [19]

B Tables

Table 6: Power Law in Firms Size, All Firms

	MLE	CDF	$\ln(Rank-0.5)$
α	1.0271	1.072	1.072
s.e.	(0.0186)	(0.00100)	(0.02150)
R^2	-	0.9957	0.9953
No. of firms	5071($\pm$655)	5113	5114
p-value	**0.2140**	-	-

Notes: This table reports the estimates of the power law exponents in firm size (total sales), using the three methods described in the text. Column (1) reports the MLE estimate given in Equation (13). Column (2) estimates the CDF of the power law specified in Equation (15). Column (3) regresses the $\ln(i-0.5)$ of the firm size distribution on the log of its size, i.e., $\ln(S_{(i)})$. The p-value quantifies the plausibility of the power law hypothesis, the critical value is 0.10 (bold if statistically significant).

Table 7: Power Law in Firms Size, Manufacturing Firms

	All firms	Exporting firms	Non-exporting firms
α	0.8711	0.8018	1.2020
s.e.	0.0453	0.0953	0.0870
n_{tail}	3066 ± 902	1956 ± 833	1306 ± 528
n	10760	6233	4527
p-value	0.00	0.00	0.01

Notes: This table reports the maximum likelihood estimates of the power law exponents in firm size (total sales), for manufacturing firms. The p-value quantifies the plausibility of the power law hypothesis, the critical value is 0.10 (bold if statistically significant). n_{tail} represents the number of observations above the estimated minimum threshold x_{min} and n is the total number of firms in each sample.

Table 8: Power Law in Firms Size, Tradeable Sectors

	All firms	Exporting firms	Non-exporting firms
α	1.0376	0.9341	1.3500
s.e.	0.0453	0.0598	0.0525
n_{tail}	3066 ± 902	2075 ± 676	1400 ± 388
n	27368	12023	14942
p-value	**0.1741**	0.00	**0.9456**

Notes: This table reports the maximum likelihood estimates of the power law exponents in firm size (total sales), for firms in tradeable sectors. The p-value quantifies the plausibility of the power law hypothesis, the critical value is 0.10 (bold if statistically significant). n_{tail} represents the number of observations above the estimated minimum threshold x_{min} and n is the total number of firms in each sample.

Table 9: Power Law in Firms Size, Exporting and Non-exporting Firms

	MLE		CDF		$ln(Rank - 0.5)$	
	(1)	(2)	(3)	(4)	(5)	(6)
	Exporters	Non-Exporters	Exporters	Non-Exporters	Exporters	Non-Exporters
α	0.9068	1.2376	0.9833	1.2568	0.9838	1.2552
s.e.	0.01653	0.0335	0.0017	0.0011	0.0264	0.0363
R^2			0.9913	0.9981	0.9906	0.9979
No. of firms	2487.4 ± 697	2644 ± 521	2778	2391	2779	2392
p-value	0.00	**0.4880**				

Notes: This table reports the estimates of the power law exponents in firm size (total sales) for exporters and non-exporters separately, using the three methods described in the text. Column (1) and (2) report the MLE estimate given in Equation (13). Column (3) and (4) estimate the CDF of the power law specified in Equation (15). Column (5) and (6) regress the $\ln(i - 0.5)$ of the firm size distribution on the log of its size, i.e., $\ln(S_{(i)})$. The p-value quantifies the plausibility of the power law hypothesis, the critical value is 0.10 (bold if statistically significant). Number of firms in Column (1) is reported with its error.

Table 10: Power Laws in Firms Size By Sector, Non-Exporting and Exporting Firms

	Exporting Firms				**Non-Exporting Firms**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sector	α	Std. Error	n_{tail}	n	α	Std. Error	n_{tail}	n
Mining and quarrying	0.8825	0.1835	84	165	0.9149	0.0763	256	500
Food products, beverages and tobacco products	1.0708	0.1560	214	766	1.3623	0.1506	250	1157
Textiles	1.8234	0.2994	342	2450	1.8927	0.3433	382	2139
Wood and Paper Products	0.7990	0.1515	285	511	1.2583	0.1536	222	478
Pharmaceuticals	0.7263	0.4985	48	76	0.7988	0.3661	36	48
Rubber and Plastic Products	0.8096	0.0778	571	1159	1.3753	0.1614	320	1163
Metals	0.8506	0.0749	416	1213	1.1279	0.0779	443	854
Computer, Electronic and Optical Products	0.7237	0.1304	82	121	1.2551	0.4960	38	52
Electrical Equipment	0.7851	0.1092	176	457	1.1417	0.1549	111	208
Machinery and Equipment	1.3576	0.1921	273	902	1.5982	0.2660	240	392
Transport Equipment	0.8766	0.1174	204	584	0.9719	0.0957	278	358
Other manufacturing	1.0656	0.0867	281	633	1.3381	0.1942	250	559
Electricity, gas, steam and air-conditioning supply	0.5208	0.6420	23	45	1.0893	0.3620	104	190
Construction	0.7013	0.1872	190	390	1.0804	0.1768	778	3351
Wholesale and Retail Trade, Repair of Motor Vehicles	0.8553	0.0398	813	2612	1.3169	0.0524	1085	6794
Transportation and Storage	0.6008	0.090	176	265	0.9466	0.1024	829	1850
Accommodation and Food Service Activities	1.0966	0.7262	31	52	1.4194	0.1874	675	1772
Publishing, Audiovisual and Broadcasting Activities	0.6196	0.2095	24	30	1.3592	0.3287	86	152
IT and Other Information Services	0.6765	0.2826	27	39	0.9777	0.1994	103	197
Legal, Accounting, Management and Other Activities	0.5884	0.2695	30	34	0.9050	0.0981	267	684
Administrative and Support Service Activities	2.4184	0.6465	29	54	1.1410	0.1704	679	1637

Notes: This table reports the maximum likelihood estimates of the power law exponents in firm size (total sales) for exporters and non-exporters separately, at sector-level. n_{tail} represents the number of observations above the estimated minimum threshold x_{min} and n is the total number of firms in each sample.

Table 11: Power Law in Firms Size, All Firms, Domestic Sales Only

	MLE	CDF	$\ln(Rank - 0.5)$
α	1.0425	1.081	1.081
s.e.	(0.01634)	(0.001)	(0.023)
R^2	-	0.9960	0.9956
No. of firms	5071(±655)	4370	4371
p-value	**0.1180**	-	-

Notes: This table reports the estimates of the power law exponents in firm size (domestic sales), using the three methods described in the text. Column (1) reports the MLE estimate given in Equation (13). Column (2) estimates the CDF of the power law specified in Equation (15). Column (3) regresses the $\ln(i - 0.5)$ of the firm size distribution on the log of its size, i.e., $\ln(S_{(i)})$. The p-value quantifies the plausibility of the power law hypothesis, the critical value is 0.10 (bold if statistically significant)

Table 12: Power Laws in Firms Size By Sector, All Sales and Domestic Sales

	All Sales			**Domestic Sales**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	α	s.e.	n_{tail}	α	s.e.	n_{tail}	n	Exports/Sales
Tradeable Sectors								
Mining and quarrying	0.8365	0.057	315	0.8351	0.062	338	666	0.2100
Food products, beverages and tobacco products	1.0681	0.1464	412	1.0816	0.1281	397	1921	0.1081
Textiles	1.8831	0.2830	518	2.4380	0.3635	643	4592	0.080
Wood and Paper Products	0.9039	0.066	490	0.9165	0.0782	472	1029	0.1025
Pharmaceuticals	0.7745	0.2719	59	0.8527	0.2754	60	125	0.0581
Rubber and Plastic Products	0.9193	0.0378	829	0.9343	0.0424	779	2418	0.1496
Metals	0.8915	0.0794	684	0.8970	0.0762	699	2162	0.1877
Computer, Electronic and Optical Products	0.7750	0.1016	105	0.7850	0.1142	100	188	0.0715
Electrical Equipment	0.8135	0.1104	264	0.8992	0.0996	254	697	0.2765
Machinery and Equipment	1.4175	0.1820	365	1.5388	0.2070	331	1416	0.2774
Transport Equipment	0.7323	0.1009	384	0.7732	0.060	394	1013	0.4508
Other manufacturing	1.1017	0.0766	435	1.1365	0.0819	451	1261	0.1318
Wholesale and Retail Trade, Repair of Motor Vehicles	1.0612	0.0291	2043	1.0748	0.0299	1981	9712	0.0931
Non-Tradeable Sectors								
Construction	0.9719	0.1105	1027	0.9861	0.1097	1010	3741	0.0168
Electricity, gas, steam and air-conditioning supply	1.0958	0.4561	100	1.1510	0.4672	103	235	0.0035
Transportation and Storage	0.7732	0.0593	400	0.8481	0.0549	995	2115	0.0112
Accommodation and Food Service Activities	1.0299	0.1718	750	1.0349	0.1634	749	1829	0.0051
Publishing, Audiovisual and Broadcasting Activities	0.8834	0.2143	91	0.8874	0.2003	90	182	0.0016
IT and Other Information Services	0.7709	0.1628	123	0.7730	0.1680	121	236	0.0047
Legal, Accounting, Management and Other Activities	0.8205	0.0964	312	0.8225	0.0924	312	729	0.0059
Administrative and Support Service Activities	1.1636	0.1738	553	1.1544	0.1854	607	1692	0.0013

Notes: This table reports the maximum likelihood estimates of the power law exponents in firm size, for total sales (Columns 1-3), and for domestic sales (Columns 4-6), at sector-level. n_{tail} represents the number of observations above the estimated minimum threshold x_{min} and n is the total number of firms in each sample. Exports/total sales ratio represents the overall sector openness.

Table 13: Summary Statistics,

	(1) No. of Firms	(2) Mean	(3) Std. Dev.	(4) Min	(5) Max
			All Firms		
Total Sales	40.077	28.799	249.145	1000	$2.0896E+7$
			Non-Exporting Firms		
	No. of Firms	Mean	Std. Dev.	Min	Max
Total Sales	27.048	$1.5828E+04$	$1.0204E+05$	0.1	$8.6252E+06$
			Exporting Firms		
	No. of Firms	Mean	Std. Dev.	Min	Max
Total Sales	13.029	$5.5727E+04$	$4.1010E+05$	0.1	$2.0896E+07$
Export Sales	14.231	$7.4014E+03$	$7.7543E+04$	0.1	$4.6835E+06$

Notes: This table presents the summary statistics for the variables used in the estimation. Sales figures are in thousand of Turkish Liras. Firms with total sales less than $TL1.000.000$ have been removed from the sample.

C Figures

Figure 2: The Analytical Power Law in the Melitz-Pareto Model: Multiple Export Markets [11]

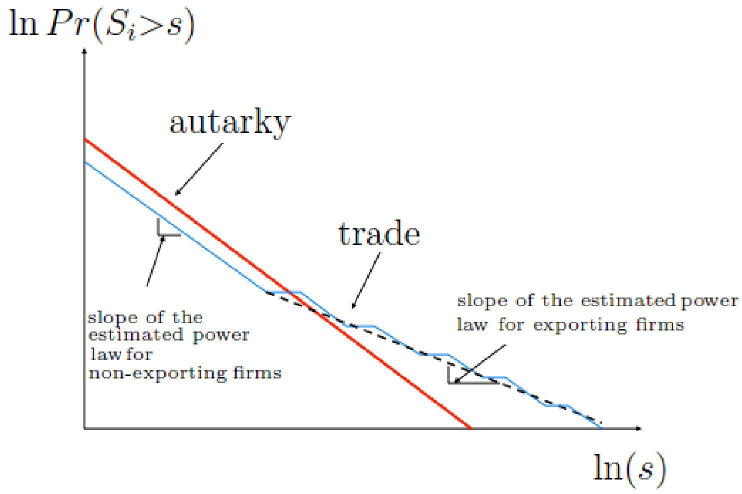


Figure 3: Partition of Firms[11]

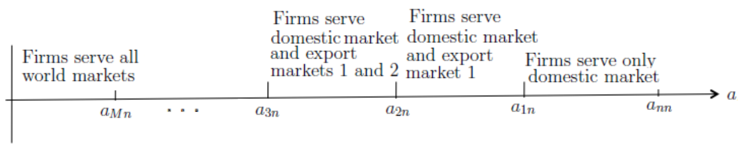


Figure 4: The Analytical Power Law in the Melitz-Pareto Model: Stochastic Export Entry Costs[11]

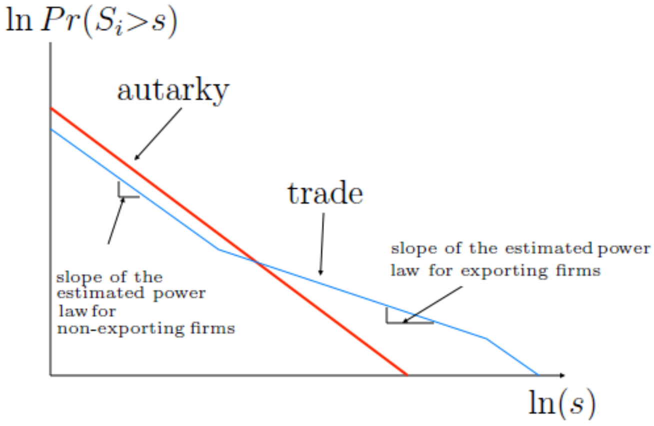
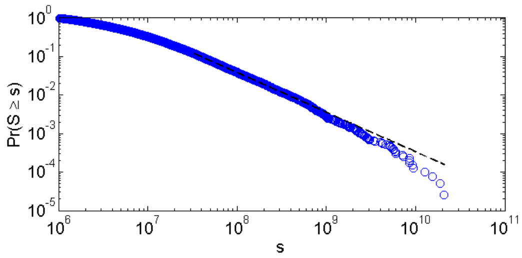
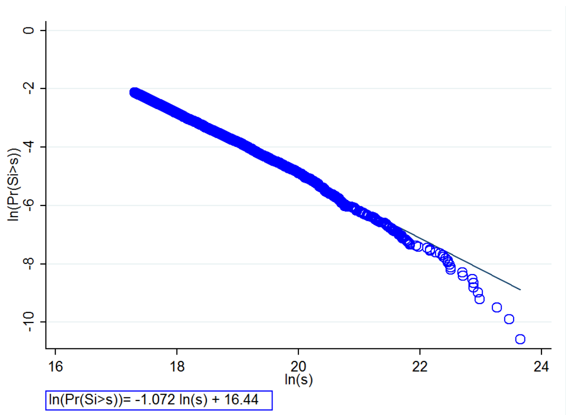


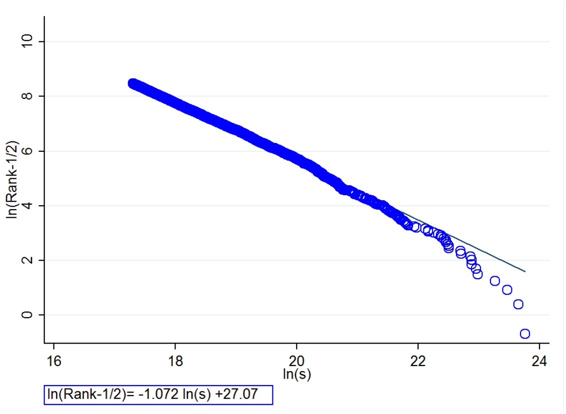
Figure 5: Power Laws in the Distribution of Firm Size, All firms



(a)



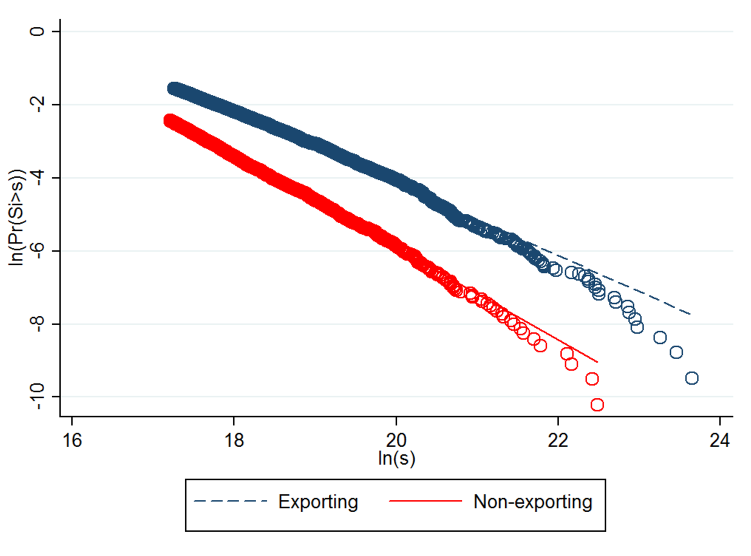
(b)



(c)

Notes: This figure reports the estimated power laws in firm size (total sales) for all firms, using the three methods described in the text: MLE (Panel (a)), the CDF (Panel (b)) and the log-rank-log-size estimation (Panel (c)).

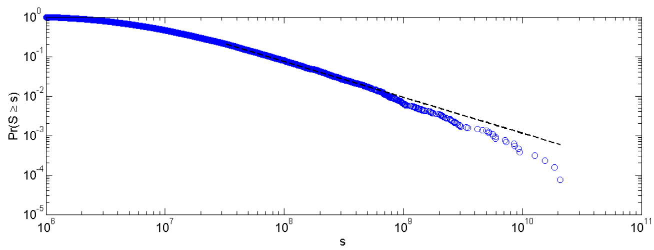
Figure 6: Exporting and Non-Exporting Firms



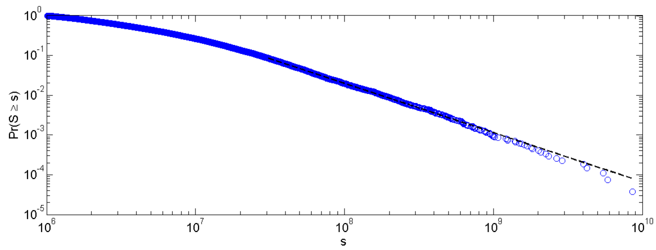
Notes: This figure reports the estimated power laws (using the CDF method) in firm size (total sales) for exporting and non-exporting firms separately.

Figure 7: Power Laws in the Distribution of Firm Size, Exporting and Non-exporting Firms

(a)

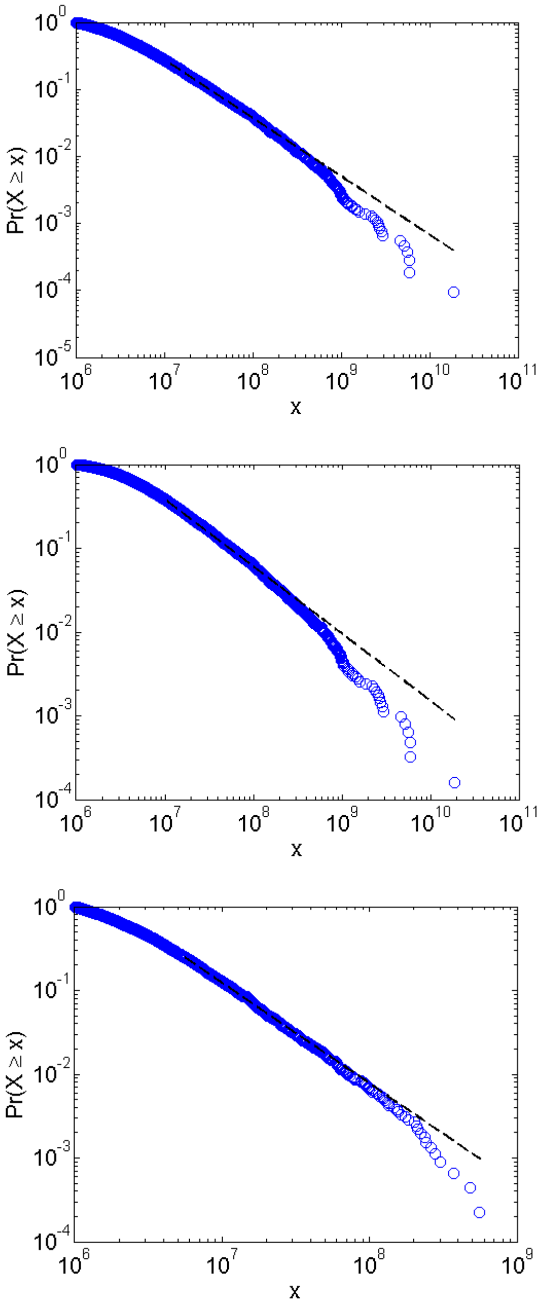


(b)



Notes: This figure reports the estimated power laws in firm size (total sales) for exporting and non-exporting firms separately. The maximum likelihood estimation is used. Exporting-firms are shown in Panel (a), non-exporting firms are shown in Panel (b).

Figure 8: Power Laws in the Distribution of Firm Size, Manufacturing Firms



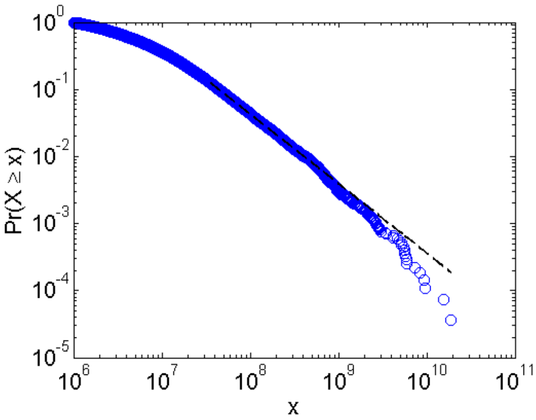
(a)

(b)

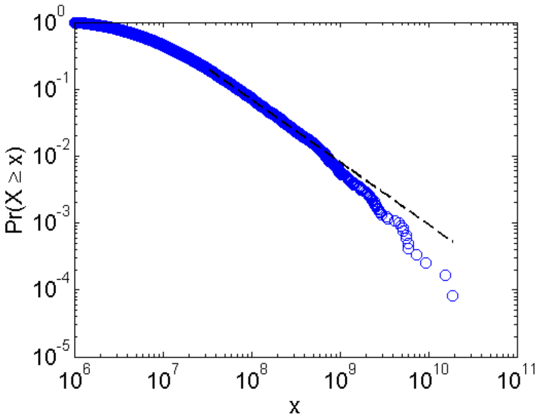
(c)

Notes: This figure reports the estimated power laws in firm size based on total sales, all manufacturing firms (Panel (a)), exporting (Panel (b)) and non-exporting manufacturing firms (Panel (c)), using the MLE method.

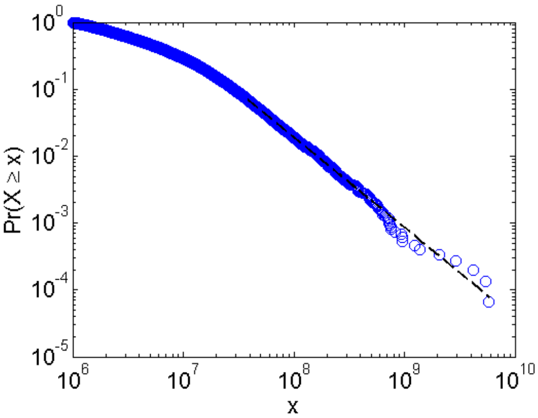
Figure 9: Power Laws in the Distribution of Firm Size, Tradeable Sectors



(a)



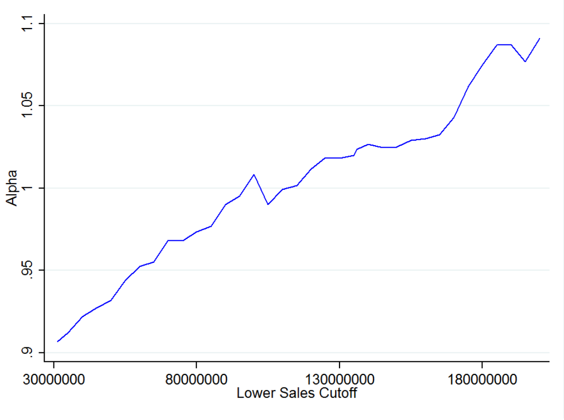
(b)



(c)

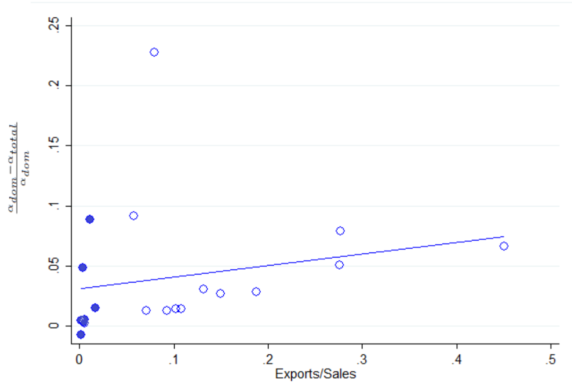
Notes: This figure reports the estimated power laws in firm size based on total sales, for firms in tradeable sectors (Panel (a)), exporting (Panel (b)) and non-exporting manufacturing firms (Panel (c)), using the MLE method.

Figure 10: Power Law Coefficient for Exporting Firms

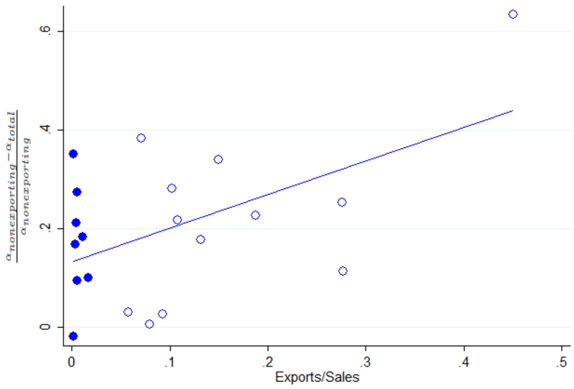


Notes: This figure depicts the estimated scaling parameter, $\hat{\alpha}$ for exporting firms, plotted against the minimum threshold.

Figure 11: Deviations in Power Law Estimates and Openness at Sector-Level



(a)



(b)

Notes: This figure plots the differences between the power law exponents at sector level against overall trade openness. In Panel (a), the percentage difference between the power law exponent estimated on domestic sales and the power law exponent estimated on total sales, i.e., $\frac{\alpha_{dom}-\alpha_{total}}{\alpha_{dom}}$ is given on the y-axis. In Panel (b), the percentage difference between the power law exponent estimated on all sales for non-exporting firms, and the power law exponent estimated on total sales of all firms, i.e., $\frac{\alpha_{nonexporting}-\alpha_{total}}{\alpha_{nonexporting}}$ is given on the y-axis Trade openness at sector level, i.e., the ratio of exports to total sales in the sector is given on the x-axis for both panels. The hollow dots represent the tradeable sectors, whereas the full ones represent the non-tradeable sectors.

Figure 12: CDF Plot for U.S. Firms (Axtell 2001, on the left) and French Firms (Di Giovanni et al. 2010, on the right), Binned Data

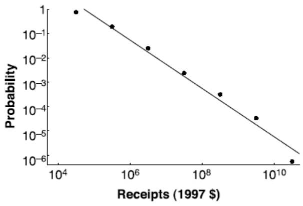


Fig. 2. Tail cumulative distribution function of U.S. firm sizes, by receipts in dollars. Data are for 1997 from the U.S. Census Bureau, tabulated in bins whose width increases in powers of 10. The solid line is the OLS regression line through the data and has slope of 0.994 (SE = 0.064; adjusted R^2 = 0.976).

Figure 4. Power Laws in the Distribution of Firm Size, All Firms